DOCUMENT RESUME

ED 110 402 SO 008 547

AUTHOR Karma, Kai
TITLE The Ability to Structure Acoustic Material as a Measure of Musical Aptitude: Part 2. Test Construction and Results. Research Bulletin 43.
INSTITUTION Helsinki Univ. (Finland). Inst. of Education.
PUB DATE Feb 75
NOTE 30p.; For a related document, see ED 092 440

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS Aptitude; *Aptitude Tests; *Comparative Testing; Fine Arts; *Music; Musical Composition; *Test Construction; Test Interpretation; Test Reliability; Test Results; Test Validity

ABSTRACT

This report is the second part of a study designed to construct a test for measuring musical aptitude of persons from various age groups. It covers the construction of the test, material, item analysis, reliability, validity, and possible future steps. The test is composed of musical recordings, determined from pilot studies, that the test groups analyzed for acoustical structure. Three versions of the test were developed to raise its reliability. Patterns of relationships instead of absolute figures are measured to show the test's validity, which is expressed in several tables in the report. Findings indicate that there are no essential differences between the total correlations and the deviations of the items in the different versions of the test. The total correlations show a relatively low but consistent positive relation. A major conclusion is that a subject's age affects his test results very little. This seems to support the theory that musical aptitude develops at an early age. See ED 092 440 for a report on the first part of this study covering the background theory and pilot studies. (ND)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED110402

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RESEARCH BULLETIN

Institute of Education
University of Helsinki

SF-00100 Helsinki 10
Finland

SP008547

No. 43

February 1975

Kai Karma

THE ABILITY TO STRUCTURE ACOUSTIC MATERIAL AS A MEASURE OF MUSICAL APTITUDE

2. Test Construction and Results

00002

AUG 1 1 19

Kai Karma

THE ABILITY TO STRUCTURE ACOUSTIC MATERIAL AS A MEASURE OF MUSICAL APTITUDE

2. Test Construction and Results

Institute of Education
University of Helsinki
1975

00003

ISBN 951-45-0529-8

00004

Preface

As this is a direct sequel to the first part of this study - Background Theory and Pilot Studies - the same persons are to be thanked for help and advice. In addition to them, I wish to thank the principal of the Music Institute of Vantaa, Olli Ruottinen, for co-operation.

It seems to be difficult to find the right place for this kind of a study in any of the faculties of the university. This being the case, I am especially grateful to the Institute of Education of the University of Helsinki for the possibility of publishing my studies.

Helsinki, January 1975
Kai Karma

Contents:

	page
1. Test construction	1
2. Material	4
3. Item analysis	5
4. Reliability	7
5. Validity	9
6. Discussion and next steps	16
References	19
Appendix: Items of B- and C-versions	

1. Test construction

When the pilot studies were completed the first test, here called version A, was composed and recorded according to the experiences gained that far. The items were directly played with an electric organ keeping the tempo subjectively relatively fast, approximately the same that was found to be suitable in the pilot studies. In addition to what has already been mentioned about the tempo - that it should be faster than the test maker feels to avoid making the test boring - there is another reason for this. It can be supposed that too slow a tempo gives the subjects a possibility to reason what the right answer is, i.e., they have time to think about several possible alternatives, eliminate the impossible ones etc., without really comprehending the holistic structure of the sounds. Because the test is made to measure an intuitive organizing ability and not reasoning ability this could be a danger to its validity.

As in the pilot studies pitch, length and intensity were used as the bases for structure forming, i.e., they vary in the items each at a time. Timbre differs from item to item to make the test more interesting. Vibrato is used in some, reverberating in some others and so on.

To avoid measuring discriminating abilities no very small differences between the sounds were used. The smallest intervals used are semitones, usually the intervals are bigger than that. In the length items the longer sounds are two times or more longer than the shorter ones. Correspondingly, the louder sounds are approximately three or four times louder than the weak ones (exact measures were not available).

The basic idea in the items has already been described in the first part of this study (Karma 1973, 16). The subjects are to divide the first part of the item into three similar parts in their minds and then decide if the second, "answer" part of the item, is similar to them. Version A consists of 40 items; in 15 of them pitch is the varying factor, in 13 and 12 items length and intensity are the bases for organizing.

Version B of the test was constructed by making an item analysis to version A and choosing the 31 best items to form the new one. Some technically imperfect items were also re-recorded.

Version C is similar to the B version except that it is not directly played but constructed by cutting and gluing a tape on which long "basic" tones had been played with an electric organ. This was done to avoid the possibility that the small imperfectnesses in the playing would lower the reliability of the test. The cutting method is relatively tedious but produces very exact results.

The exact lengths of the tones in the C version are as follows:

- In the items where pitch is the varying factor the tones follow each other immediately without a pause. In the first seven items (of which three are used as examples and for practising) the length of each sound is 0.74 seconds (14 centimetres of tape in 19 cm/sec speed), in the rest of the test the corresponding figures are 0.68 sec. (13 cm). This small shortening of the items was done to balance the subjective feeling that the tones become longer after some time of getting used to the nature of the test.

- In the items where pitch is held constant there must, naturally, be a pause between each sound. The time for this pause is taken from each sound, i.e., the sum of a sound and a pause is the same as the figures above. The length of a pause is 0.16 seconds (3 cm of tape).
- The pause between the first and the second part of an item is 3 seconds (57 cm) through the test.

The versions B and C are written in notes in the appendix.

The instruction is as follows:

The idea of dividing the first part of an item into three similar parts is presented with a couple of drawn examples, such as the following:

The subjects are shown that there is only one possible way of dividing the figure into three similar parts without leaving any figures over. When the lines showing the cut-off points have been drawn the "answer" part is compared with the first series of figures. When the drawn examples have become clear to all subjects they are told that the problems on the tape are of the same kind but the series consist of sounds instead of visual figures. The subjects are then made familiar with the test by letting them solve together the three example problems on the tape.

2. Material

The material is somewhat fragmentary owing to practical difficulties. The material is mainly obtained in connection with selecting pupils to music institutes. The institutes gave their own tests to the applicants at the same time and thus there was not very much time for tests not relevant to the selection. This is why for example the intelligence tests have been given to part of the subjects only. The different versions of the test have been used in the following connections:

-Version A has been given to the applicants for the Music Institute of Espoo in spring 1973 (N=308). Some information about intelligence, previous schooling and the tests of the institute is also available.

- Version B has been given to groups of pupils in the Music Institute of Kirkkonummi and the Pop & Jazz Institute of Oulunkylä[1)] in spring 1974. Sample sizes are 130 and 94, correspondingly. Teachers' ratings about musical aptitude and achievement are also available for part of this material.

- Version C has been given to a) the applicants for the music institutes of Espoo (N=245) and Kirkkonummi (N=44), b) the school class of the Institute of Education of the University of Helsinki (third grade, N=20), c) pupils of the Music Institute of Vantaa and some elementary school classes in Vantaa (N=133). Information about achievement, intelligence, training and the tests of the institutes is

---

[1)] The material from the Pop & Jazz Institute has been collected and processed by Irmeli Himberg, Kauko Salmi and Sampo Suihko.

also available for part of the material. C-version has been used during the spring 1974. When the figures above are summed up an overall total of 974 is attained. As to the age, the range of the subjects is from six-year-olds to adults.

3. Item analysis

There were no essential differences between the item-total correlations and the deviations of the items in the different versions of the test. Closer information is given of the version C because it is written in notes (appendix) and because of its bigger available numerus when compared to version B.

The item-total correlations show a relatively low but consistent positive relation. In this kind of a situation removing of items does not improve the reliability substantially. This being the case in all the versions of the test all items are included when reliability and validity are discussed if not especially mentioned.

The relatively low average correlations are probably most due to two reasons: first, many items have been very easy and have thus extreme p-values. These can not correlate very highly with any external variable. Second, the way of answering, the true-false format of the test, makes random guessing relatively probable which in turn lowers the correlations.

It is a common phenomenon that it is difficult to construct good items the answers to which are actually right, i.e., it is much easier to make a wrong alternative look right

Table 1. Item-total correlations (r_{it}), p-values (p) and standard deviations (s). Version C. N=309.

item	1	2	3	4	5	6	7
r_{it}	.119	.164	.155	.244	.254	.204	.266
p	.99	.95	.99	.95	.89	.97	.93
s	.08	.21	.05	.20	.30	.15	.24

item	8	9	10	11	12	13	14
r_{it}	.184	.160	.484	.333	.266	.145	.252
p	.84	.77	.72	.85	.95	.90	.55
s	.35	.41	.44	.35	.20	.29	.49

item	15	16	17	18	19	20	21
r_{it}	.183	.344	.134	.404	.256	.431	.266
p	.95	.50	.33	.76	.92	.88	.95
s	.19	.50	.47	.42	.26	.32	.20

item	22	23	24	25	26	27	28
r_{it}	.318	.193	.187	.360	.311	.293	.247
p	.93	.92	.92	.80	.90	.67	.70
s	.25	.25	.26	.39	.29	.46	.45

item	29	30	31
r_{it}	.461	.289	.322
p	.61	.65	.85
s	.48	.47	.35

than vice versa. This is also true here. The average p-value of the items to which the right answer is "yes" is .87 and the average p-value of the "no"-items is .78. This means that "no"-items discriminate better. They also seem to be better in terms of item-total correlations - if the 10 best items are chosen nine of them are "no"-items.

4. Reliability

Table 2. Reliability coefficients

	coefficient alpha	retest-reliability
version A	.66 (N=286, does not include 6 and 7 years old subjects)	.57 (correlation between versions A and B, N=37)
version B	.51 - .61 (before and after removing 10 items, N=94) .58 (.66)[1] (N=130)	.68 (N=27)
version C	.55 (.61)[1] (N=309)	

[1] The coefficients in parentheses are the reliabilities Spearman-Brown-corrected to the length of 40 items. This has been done to ease the comparison between the different versions of the test.

The basic reason for making the different versions of the test was an attempt to raise the reliability by making the tape technically better. It was thought that the small differences, e.g. in the lengths of sounds intended to be of the same length could be an important source of unreliability. However, the reliabilities could not be raised in this way which shows that the small unexactnesses in the tape are not important in general, although they may have an effect in some individual cases.

The obvious reason for the relatively low reliability of the test is, then, the true-false format of the items which makes it possible to guess right in 50 % of the cases. Oosterhof and Glassnapp (1974) have compared the reliabilities of the true-false and four-alternative multiple-choice formats empirically. According to their results the approximate reliability of .60 obtained here would be in the region of .85 - .90 if there were four alternatives to choose from in every item. Although using the multiple-choice format in this test has proved difficult when compared with tests in which the problems are presented on paper its advantages are so evident that it seems to be worth trying. The main concern is probably how this could be done without affecting the validity. One possible practical solution is presented on page 17.

Reliability coefficients provide some information about the internal construction of the test, too. Because coefficient alpha is a measure of the internal consistency of the test, and retest reliability gives information about the stability of the test over time, it would be reasonable to expect higher numerical values for the retest coefficients if there were subscales in the test, i.e., if it were not internally as consistent as the results are reproducable. Because the different coefficients are very close to each other it can be concluded that there are no clear subscales in the test.

5. Validity

In the typical case when a test is made there is no direct and reliable measure of the property aimed at; if there were one, making the test would usually be unnecessary. This makes it much more difficult to determine the validity of a test than its reliability. There are different ways of solving this problem; the strategy that was considered best in this case was to find a pattern of relations instead of a single maximized measure. This pattern of relations can be compared to the relations which are hypothesized to be present if the test measures the right theoretical concept (and if such a concept has correspondence in reality). In other words it can be said that construct validity is the main concern in this chapter. (For construct validity see, e.g., Cronbach 1966, 120; Nunnally 1967, 83). Focusing on the pattern of relations instead of the absolute figures is also reasonable here because the relatively low reliability of the test tends to cause "shrinkage" in the correlations, i.e., the level may be lower than it would be within a more reliable test although there is no reason for the relations of the correlations to be changed.

The following tables present correlations between the different varsions of the test and some other measures. Because all the measures are not available for all subjects there are empty entries in the matrices. The variables in the matrices are as follows:

Playing. The applicants for the music institutes were to play something if they had any previous experience with any instrument. The performance was rated by experienced instrument teachers. The performance was prepared in advance.
Singing. The applicants were also to prepare a little song or tune they either sang of hummed. This was rated by the same judges as above.

Tests developed at the institutes. The most common instrument in selecting pupils to music institutes in Finland is a test where the subjects are to hum or whistle a given melody and some tones reproducing the right pitch, tap given rhythms and the like. Although there is no standardized test of this kind the variation between the institutes is little.

Former training in music. Information about former training in music was given by the applicants for the institutes. Although this was given different weights by rating the effectivenesses of the various kinds of training (music classes[1] in schools, group instruction, individual teaching, etc.) this variable is probably relatively unreliable. The effects of musical/unmusical homes and the like could not be controlled. Thus this measure must be taken as a hint only.

School mark in music is in most cases an ordinary class teacher's rating of the student's achievement in the subject "music".

Sentence completion is a subtest from Heinonen's battery of factor tests of intelligence (Heinonen 1963). This test was used as an operationalization for general intelligence. There was lack of time, and this short test was considered to give information about general intelligence although more exactly it is of course a test of verbal reasoning (Heinonen's own factor analyses support this decision).

Mirror-test. This test is also taken from Heinonen's battery. It was used to validate the hypothesis made about the relation between musical and spatial ability (Karma 1973). The items consist of figures which are either similar with a model figure or mirror-images of it. The subject's task is to separate the mirror-images from the other figures.

---

[1] The term "music class" is used here to refer to classes having an additional amount of teaching in music compared to ordinary classes. The pupils are selected for these classes according to interest in music and musical aptitude.

Music institute teacher's rating of aptitude. Instrument teachers who teach their pupils individually were asked to rate the aptitude of their pupils trying to keep their judgments free from the effects of the pupils' motivation and the amount of training.

Music institute and music class teachers' ratings of achievement. These are mostly ratings of progress in instrument playing, sight singing and the like.

Table 3. Correlation matrices. A- B- and C-versions of the structuring test and some other measures. For closer information about the variables see text.

Table 3.1. Version A

1. Version A	1.	2.	3.	4.
2. Playing	.20 (99)[1)]			
3. Singing	.06 (99)	.52 (99)		
4. Tests of the institutes	.26 (99)	.41 (99)	.58 (99)	
5. Former training	.23 (99)	.41 (99)	.21 (99)	.16 (99)
6. Sentence completion	.09 (106)	---[2)]	---	---
7. Mirror-test	.33 (48)	---	---	---

Table 3.2. Version B

1. Version B	1.	2.	3.
2. Teacher's rating of aptitude	.76 (25)[1)]		
3. Teacher's rating of achievement	.15 (91)	---[2)]	
4. Former training	.13 (116)	---	.27 (116)

[1)] Because all the measures are not available for all subjects the corresponding numerus is given after every figure.

[2)] Missing information

Table 3.3. Version C

	1.	2.	3.	4.	5.	6.	7.	8.
2.	.24 (231)[1]							
3.	.12 (238)	.31 (231)						
4.	.33 (322)	.27 (231)	.51 (238)					
5.	.01 (240)	.32 (231)	.10 (238)	.07 (238)				
6.	.05 (222)	.28 (222)	.48 (222)	.34 (222)	.19 (222)			
7.	.09 (104)	-.04 (80)	.17 (83)	.23 (84)	.01 (83)	.04 (99)		
8.	.33 (89)	.10 (53)	.09 (52)	.17 (53)	-.09 (53)	.08 (51)	---[2]	
9.	.53 (54)	---	---	---	---	---	---	---

1. Version C
2. Playing
3. Singing
4. Tests of the music institutes
5. Former training
6. School mark in music
7. Sentence completion
8. Mirror-test
9. Music class teacher's rating of achievement

[1] Numerus in parentheses

[2] Missing information

00018

Table 3.4. Correlations of the A- B- and C-versions with the external criterions. A summary of tables 3.1., 3.2., and 3.3.

Criterion	Version of the test		
	A	B	C
Playing	.20 (99)[1]	---[2]	.24 (231)
Singing	.06 (99)	---	.12 (238)
Tests of the music institutes	.26 (99)	---	.33 (322)
Music institute and music class teachers' rating of achievement	---	.15 (91)	.53 (54)
Music institute teacher's rating of aptitude	---	.76 (25)	---
Former training	.23 (99)	.13 (116)	.01 (240)
School mark in music	---	---	.05 (222)
Sentence completion	.09 (106)	---	.09 (104)
Mirror-test	.33 (48)	---	.33 (89)

[1] Numerus in parentheses

[2] Missing information

The following comments may be of help when the tables are interpreted:

- When singing is rated the main sources of variance are the quality of voice and exactness of pitch. This is in line with the practical experience of the author, too. Thus, no strong relation to organizing ability is to be expected.

- The tests of the institutes are probably clearly loaded on organizing (or structuring) ability but have also a strong connection with producing capabilities. This can be supposed to lower the correlation with the structuring test. This view is supported by the fact that the tests of the institutes correlate highest (.51 - .58) with singing.

- The correlation with ratings of aptitude is probably an overestimate caused by the small numerus of this variable.[1] It is hard to make teachers estimate their pupils' aptitudes when they are used to judge achievement. As a matter of fact, a great deal of the estimates meant to be ratings of aptitude proved to be ratings of achievement when this was controlled afterwards. So only 25 ratings of aptitude are left in the tables.

- The school mark in music has quite little to do with aptitude when ordinary classes are concerned. The strongest factors forming the school marks are probaly singing and interest in music (the correlation with singing is .48).

---

[1] It should be remembered, however, that a relatively strong relation (.60) was found in the pilot studies, too (Karma 1973).

Because of the uncompleteness of the data no formal factor analysis was performed. An "armchair factor analysis"[1] seems, however, to suggest the three following factors (The analysis is mainly based on version C (table 3.3.) which is the most complete):

I. Producing, mainly singing. The correlations on which this is based are the following:

	1.	2.	3.
1. Singing			
2. Tests of the institutes	.51		
3. School mark in music	.48	.34	
4. Playing	.31	.27	.28

II. Structuring ability. This would be based on the following correlations:

	1.	2.	3.
1. The structuring test			
2. Tests of the institutes	.31		
3. Music class teacher's rating of achievement	.53	---	
4. Mirror-test	.33	.17	---

III. Former training. This would be indicated by the correlation between former training and playing (.32 in version C, and .41 in A-version, table 3.1.).

As a conclusion from the validity data it may be said that it supports the theoretical background presented in the first part of this study. A great deal of the variance in other measures of musical aptitude can be explained using the concept "ability to structure acoustic material" although it is too early to say that musical

---

[1] The term is adopted from Kerlinger (1973, 691) and refers to subjective viewing of the correlation matrix for estimating its possible factorial structure.

aptitude is the structuring ability. The supposed relation to spatial ability seems also to be present. The difference between the correlations with general intelligence (verbal reasoning) and spatial ability is exactly the same in the 1973 and 1974 material. The difference is statistically significant at 10 % level (A-version), at 5 % level (C-version) and at 2 % level in combined material. It also seems that the structuring ability depens very little on former training.

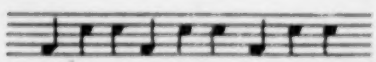
6. Discussion and next steps

The lower limit of the age of the subjects seems to be determined by their ability to read and write. Although some six years old subjects have succesfully taken the test it seems appropriate not to give the test to subjects under eight years of age. It has not been tried, however, to use the test individually. This could make it possible to test younger subjects; written answers could in this case be replaced by oral ones.

The subject's age seems to affect his results very little. Only about two points out of thirty-one was enough to balance the effect of age between eight years old and adult subjects. When compared to the effect of age on, say, intelligence tests, this is surprisingly little. This seems to support the common view that musical aptitude develops in an early age.

Using timbre for making the test more interesting and keeping it as short as possible seem to have been good decisions. Several subjects have spontaneously told that the test was nice and interesting. The younger the subjects are the more difficult and important it is to keep them motivated. Lack of motivation is thought to be able to have an effect on the validity of the test.

When the first pilot studies with the test were made it was noted that a usual multiple-choice format did not work in this kind of a test. If several alternative answers were played after the first part of an item, two problems arised: first, the subjects tended to forget the beginning of an item and, second, it was easy to guess the right answer by looking when the other subjects marked their papers. This is why the true-false format was chosen for the test. Because of the unreliability of this kind of a test a way of using the multiple-choice format without these drawbacks should, however, be developed for the future versions of the test. The solution that will be used in the next version is the following: Instead of fixing the number of similar parts in the series of sounds and making the subjects figure out what one part is like, the amount of parts varies and the task of the subjects is to determine how many similar parts the series consists of. For example, the right answer to the following item would be "three" because the series of sounds can be divided into three similar subseries:



The instruction has been a source of confusion in some cases. It is difficult to some subjects to understand the relation between the visual examples and the auditive items of the test. Thus the visual examples will be abandoned and several easy tape-recorded examples will be used to make the subjects familiar with the nature of the problems.

In addition to developing the test itself the relation of structuring ability to other abilities and personality traits will be investigated by giving the test to subjects

about whom this information is available. If there will be practical possibilities, the relation to some standardized tests of musical aptitude will also be examined.

References:

Cronbach, L.J. 1966. Essentials of Psychological Testing. New York: Harper & Row.

Heinonen, V. 1963. Älykkyystestejä lahjakkuuden eräiden alkeistekijöiden tutkimista varten. Jyväskylä: Kasvatustieteiden tutkimuskeskus.

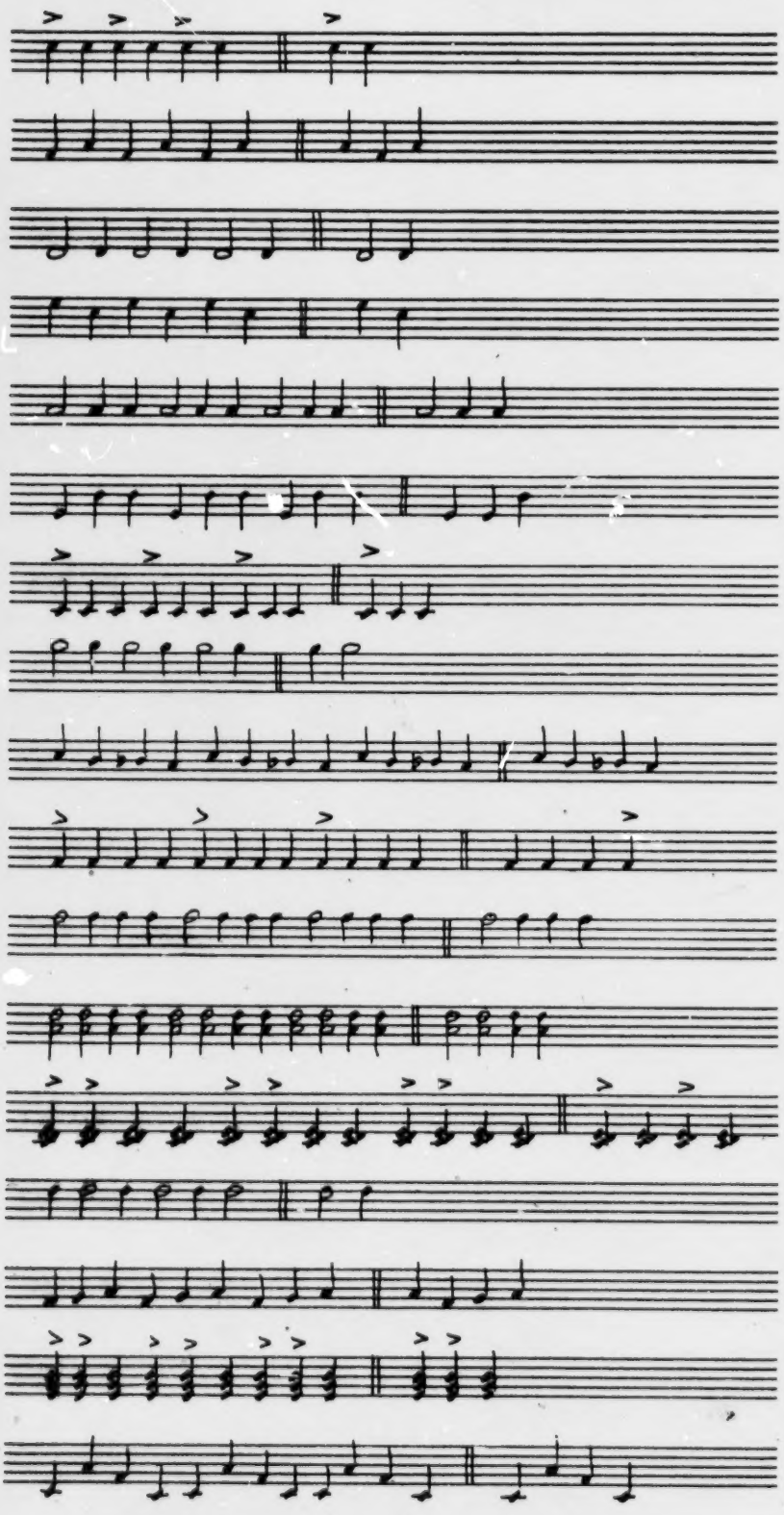
Karma, K. 1973. The Ability to Structure Acoustic Material as a Measure of Musical aptitude. 1. Background Theory and Pilot studies. Research Bulletin, Institute of Education, University of Helsinki.

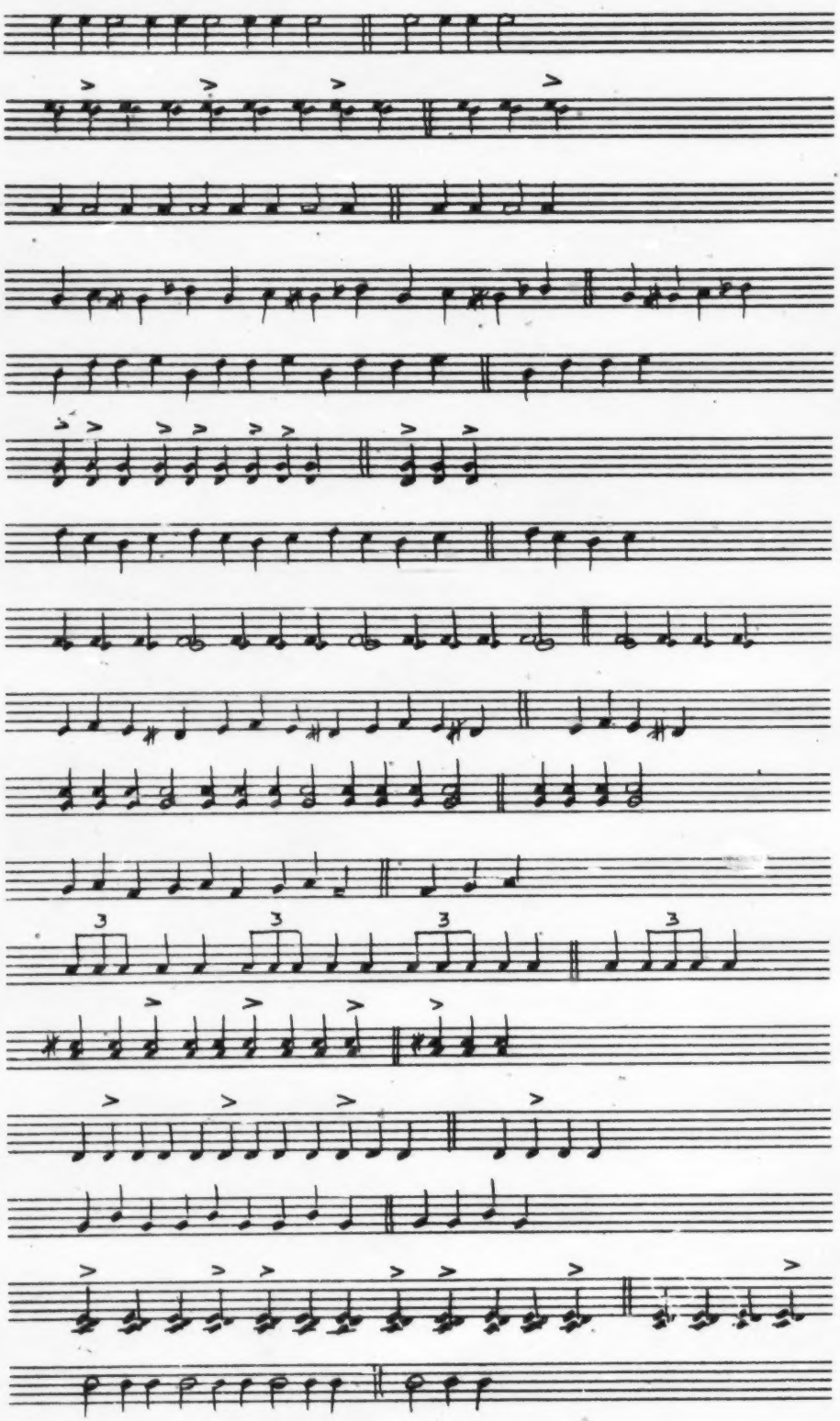
Kerlinger, F.N. 1973. Foundations of Behavioral Research. London: William Clowes & Sons Limited.

Nunnally, J.C. 1967. Psychometric Theory. New York: McGraw-Hill Book Company.

Oosterhof, A.C. and Glasnapp, D.R. 1974. Comparative Reliabilities and Difficulties of the Multiple-choice and True-false Formats. The Journal of Experimental Education, Vol. 42, N:o 3, Spring 1974.

APPENDIX. Items of B- and C-versions

item number

APPENDIX (continued). Items of B- and C-versions

item number

Previous Issues in this Series

No. 1 The Effectiveness of Punishments and the School Climate by OLLI SIPINEN. May, 1957. 10 pp. Out of print.

No. 2 The Balance of Instructional Situations by MATTI KOSKENNIEMI. Nov., 1957 (reprint 1972). 7 pp.

No. 3 On Deceit Practiced in Finnish Grammar Schools at Written Tasks by OLAVI PÄIVÄNSALO. Nov., 1957. 5 pp. Out of print.

No. 4 Comparison between Reading Comprehension and Reading Aloud by KAISA HÄLINEN. March, 1958. 9 pp. Out of print.

No. 5 Ten Test-Papers in the Swedish Language as Graded by Ten Teachers by VÄINÖ HEIKKINEN & ANITA HAVELA. Sept., 1958. 5 pp. Out of print.

No. 6 The Awarding of School Marks by TOIVO VAHERVUO. Sept., 1958. 5 pp. Out of print.

No. 7 Some Comments on the Uniformity of the 'Experimental Period' in the History of Soviet Education by MATTI KOSKENNIEMI. Feb., 1959. 18 pp. Out of print.

No. 8 Motives for the Choice of the Teaching Profession Correlated with Success in the Entrance Examination and in Student Teaching by MATTI KOSKENNIEMI. Feb., 1960. 11 pp. Out of print.

No. 9 One's Own Teacher and the Choice of Teaching Profession by MATTI KOSKENNIEMI & EIVI LIHTONEN. Sept., 1960. 14 pp. Out of print.

No. 10 The Curriculum of School Mathematics and Mathematical Performance. An Experimental Study by PAAVO MALINEN. Sept., 1961. 50 pp. Out of print.

No. 11 Prognostic Value of Some Psychological Test Results Correlated with Student Teaching Marks by MARJA-LEENA KÄYHTY. Dec., 1961. 10 pp. Out of print.

No. 12 Factors Constituting Classroom Climate by OLLI SIPINEN. March, 1963. 10 pp. Out of print.

No. 13 On the Changes in the Educational Attitudes of Young Teacher by VÄINÖ HEIKKINEN. March, 1963. 14 pp. Out of print.

No. 14 Teacher Training in Retrospection. A Preliminary Report by VÄINÖ HEIKKINEN. Feb., 1964. 21 pp. Out of print.

No. 15 Placement of Young Elementary School Teachers by MATTI KOSKENNIEMI. May, 1965 (reprint 1972). 9 pp.

No. 16 Some Problems in Combining the Individual Tutors' Marks into a Final Mark in Teaching Ability by JOHANNES ALIKOSKI. March, 1966. 25 pp. Out of print.

No. 17 The Finnish Senior Secondary Research Project I. General Presentation of the Project and Its Original Subjects by ANNA-LIISA SYSIHARJU. Sept., 1967. 31 pp.

No. 18 The Finnish Senior Secondary Research Project II. Departure from Junior Secondary and Transfer to Senior Secondary by ANNA-LIISA SYSIHARJU. Oct., 1967. 52 pp. Out of print.

No. 19 Relationships between Absences from School and Certain Factors Related to School. A Study on Elementary School Grades Three to Six by PERTTI KANSANEN. Nov., 1967. 10 pp.

No. 20 On the Anxiety Associated with School Attendance and the Grammar School Entrance Examination by EEVA PATJAS. May, 1968. 16 pp.

No. 21 School Achievement and Personality. I: Design and Hypotheses, IV: Results and Discussion by ERKKI A. NISKANEN. Oct., 1968. 51 pp.

00028

No. 22 School Achievement and Personality. II: Operations at the Variable Level by ERKKI A. NISKANEN. Oct., 1968. 124 pp.

No. 23 School Achievement and Personality. III: Operations at the Factor Level by ERKKI A. NISKANEN. Oct., 1968. 39 pp.

No. 24 The Effectiveness of Two Methods of Teaching Englich as a Foreign Language in Some Finnish Secondary Schools by DANIEL J. CASEY. Nov., 1968. 37 pp.

No. 25 The Learning of Elementary Algebra. An Empirical Investigation of the Results of Learning in a Simplified School Learning System by PAAVO MALINEN. Jan., 1969. 85 pp.

No. 26 Investigations into the Instructional Process. I. Some Methodological Problems by MATTI KOSKENNIEMI and ERKKI KOMULAINEN in collaboration with Anna-Kaarina Falck and Pentti Holopainen. Oct., 1969 (reprint 1972). 35 pp.

No. 27 Investigations into the Instructional Process. II. Objectivity of Coding in a Modified Flanders Interaction Analysis by ERKKI KOMULAINEN. Dec., 1970. 26 pp.

No. 28 Investigations into the Instructional Process. III. P-technique Treatment of Observational Data by ERKKI KOMULAINEN. Jan., 1971. 41 pp.

No. 29 Investigations into the Instructional Process. IV. Teaching as a Stochastic Process by ERKKI KOMULAINEN. Feb., 1971. 24 pp.

No. 30 Investigations into the Instructional Process. V. Experiences with the Bellack Classification System by KAI KARMA. Jan., 1972. 16 pp.

No. 31 Investigations into the Instructional Process. VI. Group Work as a Social Interaction Process. A Case Approach by PENTTI HOLOPAINEN. May, 1972. 34 pp.

No. 32 Investigations into the Instructional Process. VII. Non-verbal Behaviour in Group Work. Observational Technique and Typology of Pupils' Work by MARJA MARTIKAINEN. Sept., 1972. 18 pp.

No. 33 English School Achievements and Some Student Characteristics. I. On the Relationships of Personality and Intelligence Variables to English School Achievements by ANNA-LIISA LEINO. Oct., 1972. 50 pp.

No. 34 Investigations into the Instructional Process. VIII. On the Problems of Variable Construction from Flanders' Interaction Matrix with Special Emphasis on the Stochastic Nature of Classroom Communication by ERKKI KOMULAINEN. Feb., 1973. 78 pp.

No. 35 Zur Problematik des Unterrichts by CHR. PETROW CHRISTOW. April, 1973. 11 pp.

No. 36 Investigations into the Instructional Process. IX. Pupils' Goal-related Behavior during the Instructional Interaction by MATTI KOSKENNIEMI & PENTTI HOLOPAINEN. June, 1973. 17 pp.

No. 37 The Behaviour of Supervisors assessed by Student Teachers by SEPPO KONTIAINEN. Nov., 1973. 186 pp. Out of print.

No. 38 The Ability to Structure Acoustic Material as a Measure of Musical Aptitude. I. Background Theory and Pilot Studies by KAI KARMA. Nov., 1973. 19 pp.

No. 39 Investigations into the Instructional Process. X. Report of the DPA Helsinki, Phase One by MATTI KOSKENNIEMI & ERKKI KOMULAINEN. Feb., 1974. 35 pp.

No. 40 English School Achievements and Some Student Characteristics II. On the Relationships of Motivational and Auditory Variable to English School Achievements. Final Analyses by ANNA-LIISA LEINO. March, 1974. 48 pp.

No. 41 Foreign Language Curricuium. An Investigation into Objectives, Achievements and Student Characteristics by ANNA-LIISA LEINO. June, 1974. 27 pp.

No. 42 DPA Helsinki System for Describing Instructional Processes. Manual. Prepared and edited on behalf of the Research Team DPA Helsinki by MATTI KOSKENNIEMI, KAI KARMA and MARJA MARTIKAINEN. November, 1974. 34 pp.

Distribution by PEDUCA r.y., Fabianinkatu 28, SF - 00100 Helsinki 10

000280